y metadata, citation and similar papers at core.ac.uk
brought to you by CO
provided by Dokumenten-Publikationsserver der Humboldt-Universität zu

Wie sicher sind meine Daten auf Netzlaufwerken?

Michael Sommerfeld | Computer- und Medienservice, Systemsoftware und Kommunikation | msommerfeld@cms.hu-berlin.de

Der Beitrag richtet sich vorrangig an die Nutzer des Active Directory Windowsnetzes der HU. Er nennt einige Vorteile der durch die ZE CMS bereitgestellten Netzlaufwerke. Insbesondere werden die Möglichkeiten und Wege zur Wiederherstellung gelöschter, zerstörter, verfälschter oder fehlerhaft bearbeiteter Daten behandelt. Als Sicherungs- und Wiederherstellungssystem kam bisher und wird auch in den nächsten Jahren nur der zentrale Tivoli Storage Manager (TSM) zum Einsatz kommen. Ausgehend von der Funktionalität des TSM und der Art seiner Anwendung an der HU werden Nutzern Empfehlungen für das Verhalten bei Problemfällen gegeben.

Was zeichnet die betrachteten Netzlaufwerke aus und welcher Aspekt ihrer Datensicherheit soll hier betrachtet werden?

Dieser Beitrag betrifft nur Netzlaufwerke des Active Directory Windowsnetzes der HU, die durch die ZE CMS betreut werden. Zur Datensicherung von Netzlaufwerken, die durch andere Einrichtungen im Windowsnetz der HU bereitgestellt werden, sind die entsprechenden Verantwortlichen zu befragen.

Die durch die ZE CMS betreuten Fileserver wurden als hochverfügbare Fileserver-Cluster realisiert. Jeder Cluster besteht aus zwei bzw. drei physischen Servern, die auf die Standorte Adlershof und Mitte verteilt wurden. Alle Fileserver sind über je zwei Glasfaserpaare redundant mit dem SAN (Storage Area Network) der HU verbunden, das alle Festplatten der Netzlaufwerke bereitstellt. Das SAN der HU realisiert intern eine Spiegelung aller Platten, um weitgehende Ausfallsicherheit zu gewähren.

In diesem Artikel werden hauptsächlich die Möglichkeiten der Wiederherstellung der durch den Tivoli Storage Manager (TSM) gesicherten Daten betrachtet [1]. Die Hochverfügbarkeit wurde im vorigen Absatz nur kurz angedeutet und wird nicht weiter behandelt. Ebenso werden die geeignete Vergabe von Zugriffsrechten zu Daten und die Sicherheit der zugriffsberechtigten Accounts nicht betrachtet, obwohl das elementare Grundlagen für Datensicherheit im Internet sind.

Nutzerleitfaden

Wie stelle ich fest, welche meiner Laufwerke im Netz liegen und vom CMS betreut werden?

In der Laufwerksübersicht des Windows Explorers tauchen Zeilen folgender Struktur auf:
<freigabe>(\\<server>.<domain>) (x:)
oder
<freigabe> auf „\\<server>.<domain>“ (x:)
z.B.
Projekte (\\huuser22c.user.hu-berlin.de) (P:)
oder
Projekte auf „\\huuser22c.user.hu-berlin.de“ (P:)

Wenn als <domain> cms-, public-, user-, uva- oder uvb.hu-berlin.de steht, handelt es sich um ein zentral durch die Zentraleinrichtung des CMS bereitgestelltes Laufwerk.

Welche Richtlinien existieren an der HU zum Backup der Netzlaufwerke und welche Auswirkungen haben sie?

- Zentrale Datensicherungen werden nur durch das TSM-System realisiert.
- Platten mit Freigabenamen, die auf „...Temp“ enden (Groß-Kleinschreibung beliebig), werden nicht gesichert.
- Das TSM-System realisiert ein tägliches inkrementelles Backup aller betrachteten Netzlaufwerke – verteilt auf den Zeitraum von 18:00 bis etwa 4:30 Uhr.
- Dateien mit mehr als 10 GB werden gegenwärtig nicht gesichert.
- Es werden maximal die letzten 4 Versionen einer existierenden Datei im TSM aufbewahrt.

- Genau 4 Versionen existieren, falls an mindestens 3 Tagen innerhalb der vergangenen 30 Tage Änderungen an der Datei erfolgten. Die älteste Version stammt dann vom viertletzten Änderungs- bzw. vom Erstellungstag. Hat man unbemerkt an einem früheren Änderungstag einen Fehler bei der Veränderung der Daten begangen, so kann dieser nicht mehr durch eine Wiederherstellung aus dem TSM behoben werden! Eine eigene Generationsführung wichtiger Dateien kann das Problem etwas entschärfen. Man könnte beispielsweise nach Änderung über „abspeichern unter“ in den Namen der Datei das aktuelle Datum einfügen. Natürlich sollten dann regelmäßig alte Stände gelöscht werden, wenn man sich von der Konsistenz des aktuellen Standes überzeugt hat.
- Es existieren zwei bzw. drei Versionen, falls an genau einem Tag bzw. zwei Tagen innerhalb der vergangenen 30 Tage Änderungen an der Datei erfolgten.
- Für jede erzeugte Datei existiert nach dem ersten erfolgreichen TSM Backup eine aktuelle Sicherungsversion. Ab dem Tage nach der Erstellung der Datei sollte das der Fall sein, außer sie ist auf irgendeine Weise vom Backup ausgeschlossen. Ein Problem stellt das versehentliche Löschen einer am selben Tage erzeugten Datei dar, da noch keine Sicherung im TSM existiert und die betrachteten Netzlaufwerke nicht über einen Papierkorb verfügen, wie es bei lokalen Laufwerken der Fall ist.
- Bewusste Verfälschungen, Zerstörungen oder fehlerhafte Änderungen in einer Datei, die vor mehr als 31 Tagen erfolgten, können nie mittels TSM behoben werden, da nur noch die fehlerhafte Version im TSM existiert. Auch hier wird deutlich, dass Backup keine Archivierung ist.
- Wird eine Datei gelöscht, so behält TSM nur die aktuellste Version und markiert sie als inaktiv. Diese inaktive Kopie wird maximal 62 Tage aufbewahrt.
- Wird an Stelle einer gelöschten Datei eine gleichnamige (Groß-Kleinschreibung beliebig) erstellt, so wird die inaktive Kopie beim nächsten Backup wie eine Vorgängerversion der erstellten Datei behandelt (maximal noch 31 Tage aufgehoben), auch wenn sie es nicht ist. Dadurch können die Möglichkeiten der Wiederherstellung von Daten aus dem TSM durch den Nutzer eingeschränkt werden.
- Eine verfälschte oder verstümmelte Datei sollte nie gelöscht werden, falls die defekte Version schon durch TSM gesichert sein könnte. Ansonsten ist am Tag nach der Löschung nur noch die defekte Version im TSM vorhanden.
- Wird ein Ordner gelöscht, so wandelt TSM beim nächsten Backup die aktuelle Version in eine inaktive. Wird an Stelle eines gelöschten Ordners ein gleichnamiger (Groß-Kleinschreibung beliebig) erstellt, so wird die inaktive Kopie beim nächsten Backup gelöscht. Dies bewirkt, dass bei sogenannten ‚Point In Time' Wiederherstellungen die Ordner, die nach dem ‚Point In Time' neu angelegt und mit TSM gesichert wurden – natürlich auch alle Unterordner und enthaltene Dateien dieser –, nicht mehr als wiederherstellbare Versionen auftauchen, obwohl sie noch im TSM existieren. Ein Nutzer sollte deshalb nach Datenverlusten, die ganze Ordner betreffen, nicht versuchen, verschwundene Ordner mit gleichem Namen neu anzulegen.
- Auf dem TSM-System der HU existieren für alle Windows-Klienten (also auch für die Windows Fileserver) zentrale Dateien, die Masken für den Ausschluss bestimmter Daten von der Sicherung enthalten (siehe [2], [3]). Als Beispiele, für an der HU realisierte Ausschlüsse, seien folgende Anweisungen genannt:
  - EXCLUDE.DIR *:\...\TEMP
    alle Ordner namens „Temp“ (Groß-Kleinschreibung beliebig) werden komplett vom Backup ausgeschlossen
  - EXCLUDE *:\...\pagefile.sys
    schließt alle Dateien namens Pagefile.sys vom Backup aus
  - EXCLUDE ?:\...*.pqi
    verhindert die Sicherung aller Dateien mit der Erweiterung .pqi
- Auch lokal auf jedem Cluster-Fileserver können innerhalb der Konfigurationsdatei DSM.OPT zusätzliche EXCLUDE Anweisungen stehen und damit auf das Backup wirken. Diese Möglichkeit wird von uns aber nur zur Unterdrückung von Backup-Fehlern durch zu große oder ständig im Zugriff befindliche Dateien genutzt. Wird eine ursprünglich zu große und deshalb ausgeschlossene Datei nachträglich auf unter 10 GB verkleinert, muss wintech@cms.hu-berlin.de informiert werden, falls sie gesichert werden soll.

Welche Empfehlungen sollten bei der Nutzung der Netzlaufwerke beachtet werden?

- Es sollten nur Daten gespeichert werden, die unwiederbringlich sind – wie eigene Entwicklungen und eigene Dokumentationen. Also sollten keine zentral oder öffentlich ständig und langfristig verfügbaren Softwarepakete und Dokumentationen nochmals als persönliche Kopien vorgehalten werden.
- Wenn bestimmte Daten einer eigenen Generationsführung unterliegen, so sollte die Anzahl der Generationen geeignet beschränkt sein.
- Imagedateien ganzer Rechner oder Platten sollten nicht auf Netzlaufwerken gespeichert werden. Dafür bieten sich Wechselplatten, DVDs oder Externe Platten an.
- Archive sollten, falls keine besseren Möglichkeiten existieren, auf speziellen Netzwerkplatten zur Auswertung bereitgestellt werden. Diese Platten werden dann nicht durch TSM gesichert. Die Daten solcher Archive müssten vorab geeignet archiviert werden, da Datenverluste auf Netzwerkplatten trotz aller Sicherheitsmaßnahmen nicht ausgeschlossen sind.
- Abgeschlossene Projekte sollten geeignet archiviert/publiziert (z. B. DissOnline, Open Access, nestor, edoc) und von den Netzlaufwerken entfernt werden.
- An den Einrichtungen sollten Regelungen existieren, die den Umgang mit den Datenbeständen ausscheidender und ausgeschiedener Mitarbeiter festlegen.

Wenn Daten zerstört oder durch andere gelöscht wurden, stellt sich als erstes die Frage, wo diese Daten eigentlich gespeichert waren. Um diese Frage möglichst genau beantworten zu können, soll-

ten Sie monatlich automatisch ein Script ausführen lassen, das beispielsweise folgende Kommandofolge enthält:

```
del /f c:\baum\p_ord_old2.txt
ren c:\baum\p_ord_old.txt p_ord_old2.txt
ren c:\baum\p_ordner.txt p_ord_old.txt
p:
cd \...\ordner
dir /s > c:\baum\p_ordner.txt
```

Für \...\ordner muss der vollständige Pfad eingesetzt werden. Das Script erzeugt einen aktuellen vollständigen Überblick über die Lage aller Ordner und Dateien des Ordnerbaums ab p:\...\ordner und sichert die Strukturen der zwei Vormonate. Der Weg zur Automatisierung der Script-Abarbeitung ist vom Betriebssystem des Arbeitsplatzrechners abhängig.

Wie verhalte ich mich, wenn ich Dateien/Ordner versehentlich gelöscht oder verstümmelt habe?

- Sofort reagieren! Habe ich lokal eine aktuelle Kopie der Daten?
- Wenn nein, ist zu klären, ob die Dateien/Ordner im TSM gesichert sein könnten.
- Wenn ja, sollte ein neuer Ordner angelegt werden (z. B. \Restore), in dem die entsprechenden Daten aus dem TSM wiederhergestellt werden können. Es sollte darauf geachtet werden, dass genug freier Speicherplatz für die Wiederherstellung vorhanden ist (eventuell das Speicherlimit zeitweilig erhöhen lassen). Danach ist eine Mail an die Netzwerkverantwortlichen der Einrichtung zu schicken, die Folgendes beinhalten sollte:
  - Zeitpunkt des Datenverlustes, falls dessen Bestimmung möglich
  - möglichst genaue Quellpfad-Namen der wiederherzustellenden Daten bereitstellen
  - genauer Zeitraum, in dem der gewünschte Datenzustand verfügbar war
  - Zielpfad und -ordner der Wiederherstellung (darf nur im Zugriff der betroffenen Nutzergruppe liegen – Datenschutz; leerer Ordner ...\Restore wird empfohlen)
  - eventuelle Speicherplatzprobleme – Erhöhung des Speicherlimits
  - (Aber mindestens sollte die Mail klären, was mit welchem Stand an welchem Speicherort wiederhergestellt werden soll.)

Wie verhalte ich mich, wenn ich Daten vermisse?

- Ich versuche, sie über geeignete Suchfunktionen zu finden.
- Wenn das erfolglos war, sollten die Dateien, die die Strukturversionen des Ordnerbaums enthalten, nach den Namen durchsucht werden.
- Wenn Namen in einer Strukturdatei gefunden wurden, sollte analog zu „Wie verhalte ich mich, wenn ich Dateien/Ordner versehentlich gelöscht oder verstümmelt habe?“ verfahren werden.
- Andernfalls sollten schreibberechtigte Kollegen befragt werden.

Wie verhalte ich mich, wenn Daten offensichtlich verstümmelt sind?

- Die verstümmelten Daten dürfen solange nicht gelöscht werden, bis die Wiederherstellung der gewünschten Vorgängerversionen erfolgreich war. Die möglichen Ursachen der Verstümmelung sollten eingegrenzt werden, um Wiederholungen zu vermeiden und den ungefähren Zeitpunkt zu ermitteln.
- Analog zu „Wie verhalte ich mich, wenn ich Dateien/Ordner versehentlich gelöscht oder verstümmelt habe?“ verfahren.

Wer ist für die Sicherung und Wiederherstellung zuständig?

- Die Sicherung wird durch die Gruppen Wintech und TSM der ZE CMS organisiert.
- Die Wiederherstellung von Daten erfordert einen Auftrag an lokale Windowsnetzverantwortliche der Einrichtung.
- Diese Verantwortlichen klären den Sachverhalt und schicken danach einen entsprechenden Auftrag zur Wiederherstellung als Mail an die Gruppe Wintech.

Fazit

Die Daten der Nutzer auf den betrachteten Netzlaufwerken können bei Bedarf sicher wiederhergestellt werden, wenn die Rahmenbedingungen beachtet worden sind. Verfälschte oder fehlerhaft bearbeitete Dateien dürfen nicht gelöscht werden, bevor geeignete Versionen dieser Dateien wiederhergestellt worden sind. Das Neuanlegen gleichnamiger Ordner oder Dateien am selben Platz kann Wiederherstellungen erschweren, wenn nicht gar verhindern.
Bisher wird im Windowsnetz keine Archivierung von Daten angeboten. Wenn dazu Bedarf besteht, sollten die DV-Beauftragten der Einrichtungen informiert werden. Prinzipiell können am CMS Möglichkeiten zur Archivierung von Daten aus dem Windowsnetz in begründeten Fällen geschaffen werden.

Literatur

[1] Weickmann, Ch.: *Fileservice mit TSM*. http://www.cms.hu-berlin.de/dl/systemservice/fileservice/tsm_html

[2] Weickmann, Ch.: *Datenausschlüsse Windows*. http://www.cms.hu-berlin.de/dl/systemservice/fileservice/tsm_1_7_html, 2008

[3] Weickmann, Ch.: *Datenausschlüsse UNIX*. http://www.cms.hu-berlin.de/dl/systemservice/fileservice/tsm_1_8_html, 2008